

Efficient and feasible state tomography of quantum many-body systems

M. Ohliger^{1,2}, V. Nesme^{1,3}, and J. Eisert¹

¹ Dahlem Center for Complex Quantum Systems, Freie Universität Berlin, 14195 Berlin, Germany

² Institute for Physics and Astronomy, University of Potsdam, 14476 Potsdam, Germany and

³ Laboratoire d'Informatique de Grenoble, 38400 Saint-Martin-d'Hères, France

(Dated: February 6, 2013)

We present a novel method to perform quantum state tomography for many-particle systems which are particularly suitable for estimating states in lattice systems such as of ultra-cold atoms in optical lattices. We show that the need for measuring a tomographically complete set of observables can be overcome by letting the state evolve under some suitably chosen random circuits followed by the measurement of a single observable. We generalize known results about the approximation of unitary 2-designs, i.e., certain classes of random unitary matrices, by random quantum circuits and connect our findings to the theory of quantum compressed sensing. We show that for ultra-cold atoms in optical lattices established experimental techniques like optical super-lattices, laser speckles, and time-of-flight measurements are sufficient to perform fully certified, assumption-free tomography. This is possible without the need of addressing single sites in any step of the procedure. Combining our approach with tensor network methods – in particular the theory of matrix-product states – we identify situations where the effort of reconstruction is even constant in the number of lattice sites, allowing in principle to perform tomography on large-scale systems readily available in present experiments.

I. INTRODUCTION

Quantum state tomography is – for obvious reasons – a procedure of great importance in a large number of experiments involving quantum systems: It amounts to reconstructing an unknown quantum state entirely based on experimental data. In many situations one indeed aims at identifying what state has actually been prepared in an experiment. This seems particularly important in the context of quantum information science, where quantum state and process tomography is now routinely applied to small, precisely controlled quantum systems [1–3]. Yet, needless to say, in a number of other contexts the reliable reconstruction of quantum states is an important aim as well.

For finite-dimensional quantum systems, conventional quantum state tomography can be performed by choosing a suitable basis of $\mathcal{B}(\mathbb{C}^d)$, i.e., the operators on the d -dimensional Hilbert space of the system in question. Then, the expectation values of these d^2 observables are being measured to some required accuracy, from which one can reconstruct the unknown density matrix ρ . The same approach, however, is doomed to failure when applied to quantum many-body systems: If one has a many-body system at hand with k lattice sites of local dimension d_l , the number of necessary different measurement settings is given by $m = d_l^{2k}$, i.e., it scales exponentially with the size, rendering the treatment even of reasonably large systems impossible. Techniques of quantum compressed sensing [4–8] allow to significantly reduce the required number of measurement settings, if the state is of rank r , to $m = \Theta(rd \log^2 d)$ (where Θ denotes asymptotic equality). If $r \ll d$, this is an impressive reduction, and gives rise to feasible quantum state tomography in medium-sized quantum systems, but this number is still exponential in the number of sites. Such a scaling cannot be overcome without further restriction of the class of possible states, simply because even a pure state needs of the order of d parameters to be described. However, if the state is not only pure but also described by a generic matrix product state (MPS), the necessary number of

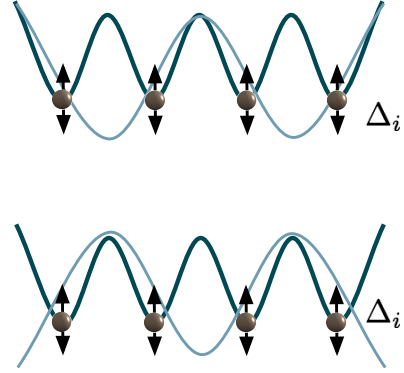


FIG. 1. Realization of a random quantum circuit by an optical superlattice. The lattice with the larger period is switched between the two depicted situations, and the lattice depth is changed locally by Δ_i which is uncorrelated between the lattice sites.

measurements only scales linearly with the system size and is even independent of it for the important special case of translationally invariant MPS. What is more, in several instances the classical procedure to reconstruct the MPS matrices from the measurement data is efficient [9].

This small number of parameters ought to make tomography an easier task, but in many practical settings involving quantum many-body systems, a serious challenge arises: In most interesting systems it is very difficult if not impossible to directly measure a full operator basis. Instead, merely measurements of some preferred observables might be readily available.

In the present work, we propose a solution to this problem. We do so by combining the action of a suitable random quantum circuit with a measurement of a very small number of different observables (even a single one can be enough). Such random circuits are just becoming a tool of great theoretical importance in several subfields of quantum information theory [10–12].

Here, we show that they also offer significant technological advantages, in that they allow for the natural implementation of *complete quantum state tomography* in systems of ultra-cold atoms using only the techniques of super-lattices and laser speckles. Both techniques have already been experimentally proven to be feasible [13, 15, 16]. Combining these new insights with the above mentioned methods of MPS tomography brings, for the first time, full tomography of many-particle systems close to experimental reach. Note that once the state is estimated, one can from this knowledge of course also compute properties such as *entanglement entropies* of systems of cold atoms [14] – but also all other properties that are a function of the quantum state.

Needless to say, it still amounts to a very difficult prescription. But while conventional measurements in ultra-cold bosonic systems amount to measuring certain correlation functions or estimates of the temperature in thermometry, say, the path described here eventually allows for the full reconstruction of an unknown state of a quantum many-body system – a very promising perspective.

The remainder of this article is organized as follows: First, we introduce in Section II tomographically complete sets of observables and a generalization of operator bases, called tight frames, and discuss their realization by means of random unitary matrices. In a next step, we show in Section III how efficient compressed sensing is possible with families of unitary matrices which form approximate unitary 2-designs [10–12, 17–19]. In Section IV, a way to efficiently realize such unitary 2-designs with the help of random quantum circuits is presented, before we discuss the application of this approach to ultra-cold atoms in optical lattices in Section V. Finally, Section VI shows how, under certain assumptions on the state, tomography is possible with a number of measurements which is linear or constant in the number of lattice sites before a conclusion is given in Section VII.

II. TOMOGRAPHY BY MEANS OF UNITARY EVOLUTION

Before we turn to the question of how tomography is possible, we provide general conditions for sets of observables to be suitable for reconstructing quantum states. We discuss why tomographic completeness as such is not sufficient and introduce generalizations of operator bases which allow for robust tomography.

A. Tomographically complete sets of observables and tight frames

We consider a quantum system consisting of k subsystems, called (lattice) sites which all have local dimension d_l . Let $\rho \in \mathcal{S}(\mathcal{H})$ be a quantum state on the Hilbert space $\mathcal{H} = \mathbb{C}^d$ with dimension $d = d_l^k$ and S the set of corresponding normalized observables, i.e., Hermitian matrices $w \in \mathcal{B}(\mathcal{H})$ with $\|w\|_2 = 1$. In this whole work, we denote by $\|\cdot\|_p$ the Schatten p -norm, where $p = 2$ and $p = \infty$ are the Frobenius norm and

operator norm respectively. Furthermore, we use the Hilbert-Schmidt scalar product defined as $(A, B) = \text{Tr}(A^\dagger B)$ and the projection on the subspace spanned by some $w \in S$ which is defined as

$$\mathcal{P}_w : \rho \mapsto (w, \rho)w. \quad (1)$$

Noting that a measurement of the expectation value of w corresponds to determining (w, ρ) , we define the total sampling operator

$$\mathcal{W}_d = d^2 \int d\mu(w) \mathcal{P}_w \in \mathcal{B}(\mathcal{B}(\mathcal{H})), \quad (2)$$

where μ is a probability measure on S .

A finite set of observables is said to be tomographically complete if any two different quantum states have distinct expectation values for some observable: This implies that one can theoretically reconstruct the state if one knows all expectation values. If \mathcal{W}_d has full rank, i.e. $\text{rank}(\mathcal{W}_d) = d^2$, the state ρ can be obtained from $\mathcal{W}_d(\rho)$ by matrix inversion if issues of statistical errors and numerical imprecisions are neglected. For general probability measures we make use of the subsequent definition.

Definition 1 (Tomographic completeness). *The measure μ is said to be tomographically complete if \mathcal{W}_d , as defined in Eq. (2), is full rank.*

In practice, not every tomographically complete measure on the observables is necessarily useful: The observables, viewed as vectors in \mathbb{R}^{d^2} , should not be too unevenly distributed over the sphere. That is to say, if the ratio between the largest and smallest eigenvalues of \mathcal{W}_d is large, small errors in the expectation values can lead to large errors in the reconstructed state. The ideal situation is that of a tight frame, also known as spherical 1-design:

Definition 2 (Tight frame). *A probability measure μ on the set of 2-norm normalized Hermitian matrices S is called a tight frame if $\mathcal{W}_d = \mathbb{1}$ with \mathcal{W}_d given by Eq. (2).*

Examples for tight frames include any operator basis and the rotationally invariant measure on the 2-norm sphere S [8]. When observables are taken from tight frames, the reconstruction problem is well conditioned and small errors in the expectation values only lead to a small error in the reconstructed state [8, 20].

B. Evolution of observables

A tomographically complete set of observables must contain at least d^2 observables, which might be difficult to measure directly. We introduce a way to do tomography by performing suitably random unitaries, followed by the estimation of the expectation value for a single observable. A different but related approach has been employed in Ref. [21] to perform high-fidelity quantum state reconstruction in situations where the knowledge about the state is not tomographically

complete. In the present work, the time evolution is a tool to obtain knowledge about the quantum state of the system leaving the question of determining Hamiltonians aside [22].

Switching to the Heisenberg picture, the outlined procedure amounts to measuring the time-evolved observables. Simple as this idea is, it allows the economical reconstruction of unknown quantum states, as it turns out:

Definition 3 (Induced observables). *To a measure $\tilde{\mu}$ on the special unitary group $SU(d)$ and an observable w_0 we associate the following induced measure on S :*

$$\mu = \left(1 - \frac{1}{d^2}\right) \tilde{\mu} \circ f^{-1} + \frac{1}{d^2} \delta_{\mathbb{1}/\sqrt{d}} \quad (3)$$

where $f : U(d) \rightarrow S$ is defined by $f(U) = U^\dagger w_0 U$, and δ_x denotes the Dirac in x . Notice that if $\tilde{\mu}$ is a probability measure, then so is μ .

An important insight is provided by the following observation.

Theorem 1 (Tight frame induced by random unitary). *Let w be a traceless, normalized observable and μ_H be the Haar measure on $SU(d)$. The measure induced on S is a tight frame.*

Proof: To show this observation, we use (3) to calculate the sampling operator (2) for a state ρ with $\text{Tr}(\rho) = 1$:

$$\mathcal{W}_{d,H}(\rho) = (d^2 - 1) \int d\mu_H(U) (U^\dagger w_0 U, \rho) U^\dagger w_0 U + \frac{\mathbb{1}}{d}. \quad (4)$$

As w is diagonalized by a unitary matrix, we can assume it to be diagonal and obtain

$$(\mathcal{W}_{d,H}(\rho))_{i,j} = (d^2 - 1) \sum_{k,l,m,n} \mathbb{E}_{U \sim \mu_H} [U_{k,l} U_{m,j} \bar{U}_{k,n} \bar{U}_{m,i}] \times w_{k,k} w_{m,m} \rho_{l,n} + \frac{\delta_{i,j}}{d} \quad (5)$$

where $U \sim \mu_H$ indicates that U is distributed according to the Haar measure. The occurring expectation values can be obtained from Ref. [23]. We first consider the off-diagonal elements, i.e., the ones with $i \neq j$. Here, the expectation value vanishes unless $l = i$ and $n = j$, in which case we get

$$\mathbb{E}_{U \sim \mu_H} [U_{k,i} U_{m,j} \bar{U}_{k,j} \bar{U}_{m,i}] = \frac{1}{d(d^2 - 1)} (\delta_{k,m}(n - 1) + (1 - \delta_{k,m})(-1)). \quad (6)$$

We now turn to the diagonal elements and note that if $i = j$, one needs $l = n$ to get a non-vanishing expectation value. We consider two cases separately: If $m = k$ we get

$$\mathbb{E}_{U \sim \mu_H} [U_{k,l} U_{k,i} \bar{U}_{k,l} \bar{U}_{k,i}] = \frac{1}{d(d^2 - 1)} (\delta_{i,l} 2(n - 1) + (1 - \delta_{i,l})(n - 1)), \quad (7)$$

while for $m \neq k$ we obtain

$$\mathbb{E}_{U \sim \mu_H} [U_{k,l} U_{m,i} \bar{U}_{k,l} \bar{U}_{m,i}] = \frac{1}{d(d^2 - 1)} (\delta_{i,l}(n - 1) + (1 - \delta_{i,l})n). \quad (8)$$

Inserting now (6), (7), and (8) into (4) and using $\sum_i \rho_{i,i} = 1$, $\sum_i w_{i,i} = 0$, and $\sum_i |w_{i,i}|^2 = 1$, we get $\mathcal{W}_{d,H}(\rho) = \rho$ which concludes the proof. ■

C. Tight frames under physical restrictions

In many situations of interest, the quantum state is not completely arbitrary but satisfies some additional conditions. In experiments with ultra-cold atoms, for example, the total particle number is conserved and super-selection rules forbid superpositions of states belonging to different eigenvalues of the corresponding operator. Assume that the quantum state ρ acts only in some subspace and denote the corresponding projection super-operator as \mathcal{P}_N . An important example is given by eigenspaces of an operator \hat{N} describing a conserved quantity, i.e. one with commuting with ρ . In this case of a quantum state confined to a subspace, Definition 2 can be relaxed to

$$\mathcal{W}_{d_N} \circ \mathcal{P}_N = \mathcal{P}_N \quad (9)$$

where d_N denotes the dimension of the subspace corresponding to the eigenvalue N of \hat{N} . If one aims at realizing such a restricted tight frame by the means of Theorem 1, one can replace the group $SU(d)$ by

$$SU_{\hat{N}}(d) = \{U \in SU(d) : [U, \hat{N}] = 0\}. \quad (10)$$

As this group is compact, there exists a unique Haar probability measure on it. We adapt Definition 3 to this situation.

Definition 4 (Induced observable on subspace). *Let $\hat{N} \in \mathcal{B}(\mathcal{H})$ and S_N its d_N dimensional eigenspace corresponding to the eigenvalue N . To a measure $\tilde{\mu}$ on the group $SU_{\hat{N}}(d)$ and an observable $w_0 \in S_N$ with $\|w_0\|_2 = 1$, we associate the following measure on S_N :*

$$\mu = \left(1 - \frac{1}{d_N^2}\right) \tilde{\mu} \circ f^{-1} + \frac{1}{d_N^2} \delta_{\mathbb{1}/\sqrt{d_N}} \quad (11)$$

where $f(U) = U^\dagger w_0 U$.

The matrices ρ , w_0 , and all $V \in SU_{\hat{N}}(d)$ are block diagonal with respect to the eigenbasis of \hat{N} . We consider the block corresponding to the eigenvalue N . Invariance of the Haar measure on $SU_{\hat{N}}$ implies that if $V \sim \mu_{H,SU_{\hat{N}}(d)}$, then $\mathcal{P}_N(V) \sim \mu_{H,SU(d_N)}$. Thus, one can apply the proof of Theorem 1 to this block and obtain Eq. (9).

III. UNITARY t -DESIGNS

The effort to implement random unitaries drawn from the Haar measure scales exponentially in the number of lattice

sites k , making an implementation both theoretically inefficient and practically unfeasible. However, this problem can be circumvented by replacing the Haar measure by a unitary 2-design which is much easier to sample from as we will see later on. Unitary t -designs behave like the Haar measure in specific situations [12, 18, 19]. The definition most suited to our problem is the following: Let ν be a probability measure on $SU(d)$. We define two channels on $\mathcal{B}(\mathcal{H}^{\otimes t})$:

$$\mathcal{G}_{t,\nu}(\rho) = \mathbb{E}_{U \sim \nu} [U^{\otimes t} \rho (U^\dagger)^{\otimes t}] \quad (12)$$

and \mathcal{G}_{t,μ_H} , where μ_H is the Haar measure. We say that ν is a unitary t -design if $\mathcal{G}_\nu = \mathcal{G}_H$. We say that ν is an ε -approximate t -design if

$$\|\mathcal{G}_{t,\nu} - \mathcal{G}_{t,H}\| \leq \varepsilon \quad (13)$$

where $\|\cdot\|$ denotes the superoperator $2 \rightarrow 2$ -norm which is defined as

$$\|\mathcal{O}\| = \sup_{X, \|X\|_2=1} \|\mathcal{O}(X)\|_2. \quad (14)$$

This superoperator norm is equal to the Schatten ∞ -norm when the channel is seen as a mere linear operator acting on the real vector space of Hermitian matrices. In the remainder we only consider the case $t = 2$ and drop the index t for simplicity.

A. Tight frames from unitary 2-designs

For our purpose, i.e., replacing the Haar measure random unitary in Theorem 1, we need an approximate unitary 2-design. This is the case because when Haar measure induces a tight frame, c.f. Eq. (4), both U and U^\dagger appear twice.

Theorem 2 (Tight frames induced by unitary 2-designs). *Let w_0 be a traceless, 2-norm normalized observable and ν be an ε -approximate unitary 2-design. Then the sampling operator $\mathcal{W}_{d,\nu}$ corresponding to the induced measure fulfills*

$$\|\mathcal{W}_{d,\nu} - \mathbb{1}\| \leq \sqrt{d}(d^2 - 1)\varepsilon. \quad (15)$$

Proof: We have

$$\begin{aligned} \|\mathcal{W}_{d,\nu} - \mathbb{1}\| &= \|\mathcal{W}_{d,\nu} - \mathcal{W}_{d,H}\| \\ &= \sup_{X, \|X\|_2=1} \|\mathcal{W}_{d,\nu}(X) - \mathcal{W}_{d,H}(X)\|_2. \end{aligned} \quad (16)$$

We note that

$$\begin{aligned} \mathcal{W}_{d,\nu}(X) &= (d^2 - 1) \mathbb{E}_{U \sim \nu} [(U^\dagger w_0 U, X) U^\dagger w_0 U] + \frac{1}{d} \text{Tr} X \\ &= (d^2 - 1) \sqrt{d} \text{Tr}_1 \left(\mathbb{E}_{U \sim \nu} [(U^\dagger \otimes U^\dagger) \right. \\ &\quad \left. \times (w_0 \otimes w_0)(U \otimes U)] \left(X \otimes \frac{\mathbb{1}}{\sqrt{d}} \right) \right) + \frac{1}{d} \text{Tr} X, \end{aligned} \quad (17)$$

where Tr_1 denotes the partial trace with respect to the first of the two subsystems of equal dimension. Note that we have extended the definition of $\mathcal{W}_{d,\nu}$ to operators with non-unit trace. This relation yields (15) after inserting it into Eq. (16) and applying (13). ■

The same argument holds also for restricted tight frames as defined in (9): Let ν be a distribution on $SU_{\hat{N}}(d)$ such that if $U \sim \nu$, then $\mathcal{P}_N(U)$ is drawn from an ε -approximate 2-design on $SU(d_N)$. This implies

$$\|\mathcal{W}_{d_N,\nu} \mathcal{P}_N - \mathcal{P}_N\| \leq \sqrt{d_N}(d_N^2 - 1)\varepsilon, \quad (18)$$

as follows from the application of the above proof to the block in U corresponding to the eigenvalue N of \hat{N} .

B. Compressed sensing

The technique of compressed sensing allows to reduce the number of measurements which are needed to reconstruct a quantum state from $\Theta(d^2)$ to $\Theta(d \text{polylog}(d))$ if the rank of the state does not increase with d . To perform this method, one has to choose $m = \Theta(d \text{polylog}(d))$ observables w_1, \dots, w_m randomly from the tight frame according to the corresponding probability measure and determine their expectation value (w_i, ρ) by measurement. Then, one can efficiently solve the optimization problem

$$\min_{\sigma} \|\sigma\|_1 \text{ s.t. } \forall i = 1, \dots, m : (w_i, \sigma) = (w_i, \rho). \quad (19)$$

The theory has been developed for observables forming operator bases in Refs. [5, 6] and extended to tight frames in Ref. [8]. There, it was also shown that the tight frame condition may be violated and compressed sensing is still possible if

$$\|\mathcal{W}_d - \mathbb{1}\| \leq \frac{1}{8\sqrt{r}}, \quad (20)$$

where r is the rank of the state we want to reconstruct. Not all tight frames are equally suited for compressed sensing as can be seen with a simple example: Let w_1, \dots, w_{d^2} be the elements of an orthonormal operator bases of $\mathcal{B}(\mathbb{C}^d)$ where w_1 is a rank-one projector and $\rho = w_1$. In this case, one has to measure of the order of d^2 observables before one “hits” w_1 and gets any information on the system. In Refs. [5, 6, 8], it has been shown that this problem cannot occur if all observables fulfill the so-called “Fourier type incoherence condition” which reads

$$\mathbb{P}_{w \sim \mu} \left(\|w\|_\infty^2 > \frac{\lambda}{d} \right) = 0, \quad (21)$$

where λ must fulfill $\lambda = O(\text{polylog}(d))$. Note that statements like (21) make only sense when considering families of tight frames with growing dimension d . As we are mainly interested in the asymptotic efficiency of our scheme, we restrict ourselves to the scaling behavior and omit explicit prefactors. We now give a condition under which Eq. (21) is fulfilled in the situation of interest.

Theorem 3 (Compressed sensing with induced observables). *Let w_0 be a traceless, normalized observable fulfilling $\|w_0\|_\infty^2 \leq \lambda/d$ with $\lambda = O(\text{polylog}(d))$, and let ν be a $1/(8\sqrt{rd}(d^2 - 1))$ -approximate 2-design. The induced tight frame fulfills (21), which implies that it allows for compressed sensing.*

Proof: Since

$$\|U^\dagger w_0 U\|_\infty^2 = \|w_0\|_\infty^2 \quad (22)$$

and $\|1/\sqrt{d}\|_\infty^2 = 1/d$, condition (21) is fulfilled and from (15) it follows that (20) is satisfied which proves that compressed sensing is possible. ■

Theorem 3 holds especially in the important case of an observable which acts non-trivially only on a few number of lattice sites because here it is of the form

$$w_0 = v \otimes \frac{\mathbb{1}_{d_l^{k-m}}}{(d_l^{k-m})^{1/2}} \quad (23)$$

for some normalized, traceless observable v and some small constant m .

IV. APPROXIMATION OF UNITARY 2-DESIGNS BY RANDOM QUANTUM CIRCUITS

In the previous section, we have shown that unitary 2-designs can be used to realize tight frames. We now show how they can be approximated by parallel random circuits and generalize the results of Refs. [10, 11, 17] to show the following:

Theorem 4 (Random circuits). *Assume k to be even. Consider a parallel random circuit where in each step either $U_{1,2} \otimes U_{3,4} \otimes \dots \otimes U_{k-1,k}$ or $U_{2,3} \otimes U_{4,5} \otimes \dots \otimes U_{k-2,k-1}$ is performed with probability $1/2$ where $U_{i,i+1}$ acts on the neighboring sites i and $i+1$. If the nearest-neighbor unitaries are drawn, in each step independently, from a probability measure ν_2 which is universal, as defined below, there exists a constant C (depending on the local dimension) such that the random circuit forms an ε -approximate unitary 2-design after $n = C \log(1/\varepsilon)k \log k$ steps.*

A finite set of nearest-neighbor unitary quantum gates is called universal if they generate a dense subgroup of $SU(d_l^2)$. For an arbitrary probability measure on $SU(d_l^2)$, the notion of universality can be generalized, according to Ref. [17]:

Definition 5 (Universality). *We say that μ is universal if for any open ball S there exists $l > 0$ such that S has a nonzero weight for the l -fold convolution product of μ .*

Proof: The proof of Theorem 4 is an extension of that of similar results in Refs. [10, 11]. Readers mostly interested in the application to optical lattice systems can safely skip to the next section.

In Ref. [11], it is shown that parallel random circuits with periodic boundary conditions form ε -approximate 2-designs after $n = C \log(1/\varepsilon)k$ steps if the unitaries are drawn from the Haar measure on $SU(d_l^2)$.

We now proceed in three steps: First, we show that the nearest-neighbor unitaries can be drawn from an approximate 2-design on $SU(d_l^2)$ instead from the Haar measure. We denote the measure corresponding to a single step of the random circuit by ν_k and define the linear operator G_{ν_k} by

$$G_{\nu_k} = \int d\nu_k(U) U^{\otimes 2} \otimes \bar{U}^{\otimes 2}. \quad (24)$$

This operator can be decomposed as $G_{\nu_k} = (M_e + M_o)/2$ with

$$M_e = P_{1,2} \otimes P_{3,4} \otimes \dots \otimes P_{k-1,k}, \quad (25)$$

$$M_o = P_{k,1} \otimes P_{2,3} \otimes \dots \otimes P_{k-2,k-1} \quad (26)$$

where

$$P_{i,j} = \int d\nu_2(U_{i,j}) U_{i,j}^{\otimes 2} \otimes \bar{U}_{i,j}^{\otimes 2}. \quad (27)$$

We have to bound

$$\|G_{\nu_k}^n - G_H\|_\infty = \lambda_2(G_{\nu_k})^n \quad (28)$$

where n is the depth of the circuit and λ_2 denotes the second largest eigenvalue. In Ref. [11], it is shown that if the nearest-neighbor unitaries are drawn from the Haar measure, there exists a constant $\Lambda > 0$ such that the corresponding operator \tilde{G}_{ν_k} fulfills $\lambda_2(\tilde{G}_{\nu_k}) \leq 1 - \Lambda$. Using now the fact, that ν_2 is a δ -approximate 2-design, we get

$$\lambda_2(G_{\nu_k}) \leq \lambda_2(\tilde{G}_{\nu_k}) + k\|G_{\nu_2} - G_{H_2}\|_\infty \leq 1 + k\delta - \Lambda. \quad (29)$$

For the right-hand side to be smaller than one, which is necessary and sufficient for an exponentially fast convergence, one has to choose $\delta = O(1/k)$. To realize this *local* approximate 2-design, we use a result from Ref. [17] which states that one needs to draw only $s = O(\log(1/\delta))$ gates from an arbitrary universal gate set to achieve this. We pick s to be a power of 2, which is surely always possible. Thus, we have a circuit with depth

$$n_\varepsilon = O(ns) = O(\log(1/\varepsilon)k \log k) \quad (30)$$

where the random choice between M_e and M_o is not made in every step but in blocks of s steps which corresponds to the operator $(M_e^s + M_o^s)/2$ while s steps of the actual quantum circuit performed are described by $((M_e + M_o)/2)^s$. They both have the same fixed point. What is more, Theorem 5 holds true, see below. This means, in particular, that the convergence of the actual circuit cannot be slower and (30) holds.

The last thing needed to obtain Theorem 4 is to switch to open boundary conditions, i.e., remove the first tensor factor in Eq. (26). As this does not affect the fixed point and the operator norm difference is on the order of δ , only the prefactor is changed slightly. We note that the prefactor depends on the actual choice of ν_2 .

The conditions for Theorem 3 to apply are fulfilled if $1/\varepsilon = O(d^{5/2})$. Using this in Eq. (30), we get

$$n_T = O(k^2 \log k). \quad (31)$$

This means compressed sensing is possible with a single traceless observable and a parallel random quantum circuit with a depth given by Eq. (31).

We did not explicitly discuss the case of restricted tight frames because it can be, again treated by block-decomposing all matrices according to the spectral decomposition of the operator \hat{N} describing the symmetry, c.f. Eq. (10). Thus, with a parallel random quantum circuit as in Theorem 4 with unitaries which are universal for $SU_{\hat{N}}(d_l^2)$ with a depth as in Eq. (31), one can perform compressed sensing for states within some eigenspace of \hat{N} . ■

Theorem 5 (Mixing properties of circuits). *Let s be a power of 2, then*

$$\lambda_2 \left(\left(\frac{M_e^2 + M_o^2}{2} \right)^s \right) \leq \lambda_2 \left(\frac{M_e^{2s} + M_o^{2s}}{2} \right), \quad (32)$$

Proof: We now prove the validity of Eq. (32) for s being a power of 2, i.e. $s = 2^j$. Let A and B be two Hermitian matrices. From $0 \leq (A - B)^2$ it follows that

$$\left(\frac{A + B}{2} \right)^2 \leq \frac{A^2 + B^2}{2} \quad (33)$$

and hence also, from the monotonicity of eigenvalues [28],

$$\lambda_2 \left(\left(\frac{A + B}{2} \right)^2 \right) \leq \lambda_2 \left(\frac{A^2 + B^2}{2} \right). \quad (34)$$

We also need the fact that, for all pairs of positive Hermitian matrices A and B ,

$$\lambda_2(B) \geq \lambda_2(A) \Rightarrow \lambda_2(B^i) \geq \lambda_2(A^i), \quad (35)$$

which can be easily seen by diagonalizing A and B . Note that M_e^2 and M_o^2 are positive operators. For any natural i_1, i_2 one finds that

$$\begin{aligned} \lambda_2 \left[\left(\frac{M_e^{2^{i_1}} + M_o^{2^{i_1}}}{2} \right)^{2^{i_2}} \right] &= \lambda_2 \left[\left(\left(\frac{M_e^{2^{i_1}} + M_o^{2^{i_1}}}{2} \right)^2 \right)^{2^{i_2-1}} \right] \\ &\leq \lambda_2 \left[\left(\frac{M_e^{2^{i_1+1}} + M_o^{2^{i_1+1}}}{2} \right)^{2^{i_2-1}} \right], \end{aligned} \quad (36)$$

where we have employed (35) and (34) for $A = M_e^{2^{i_1}}$ and $B = M_o^{2^{i_1}}$. By a repeated application of (36), starting from $i_1 = 1$ and $i_2 = j$, until $i_1 = j$ and $i_2 = j$, we finally get

$$\lambda_2 \left(\left(\frac{M_e^2 + M_o^2}{2} \right)^{2^j} \right) \leq \lambda_2 \left(\frac{M_o^{2^{j+1}} + M_e^{2^{j+1}}}{2} \right), \quad (37)$$

which is the statement to be shown. ■

V. OPTICAL LATTICE SYSTEMS

We now introduce a novel method for quantum state tomography for systems of ultra-cold atoms in optical lattices that does *not* require any local addressing of single sites. In fact, each of the steps necessary has already been demonstrated experimentally. The idea is to make use of appropriate randomness and natural time evolution, suitable exploiting optical superlattices, such that time-of-flight images give rise to complete tomographic information – quite an interesting and promising perspective. Using these experimental tools, the above mathematical methods become applicable. It should be clear, as pointed out before, that such a prescription also allows for detecting entanglement entropies [14, 29] in systems of cold atoms.

Ultra-cold atoms in optical lattices form some of the cleanest quantum many-particle systems available for experiments and allow for the realization of well-known effects from condensed matter. For example, both the bosonic super-fluid Mott-insulator transition and the Mott state of fermions where observed by changing the intensity of the laser forming the lattice [24, 25]. Such systems also have the potential of functioning as quantum simulators which means, they allow to simulate systems from other branches of physics like the notoriously difficult quantum chromodynamics (QCD) [26].

A. Time-of-flight measurements

Even though measurements with spatial resolution have been demonstrated in recent experiments [16], the standard technique is still provided by time-of-flight absorption imaging. Here, the lattice and the confining trap are instantaneously switched off and, after some time during which the atoms expand approximately without interaction, an absorption image is taken [24, 30]. As the distance that the atoms fly during the expansion is proportional to the initial momentum, this procedure amounts to a measurement of the density in momentum space. We restrict ourselves to the bosonic case, while noting that fermions can be treated in a completely analogous way, and expand the field operators of the one-dimensional bosonic field as

$$\hat{\Psi}(x) = \sum_{j=1}^{\infty} \sum_{s=1}^k W^{(j)}(x - x_s) \hat{b}_s^{(j)} \quad (38)$$

where $W^{(j)}$ is the Wannier function of the j -th band, x_s is the position, and $\hat{b}_s^{(j)}$ the corresponding annihilation operator at site s . If the lattice is sufficiently deep, all bands but the lowest one can be neglected, and we drop the upper index in (38). The momentum-space distribution, which is measured in the time-of-flight experiment, is given by $n(p) = |\tilde{W}(p)|^2 S(p)$ where \tilde{W} is the Fourier transform of the Wannier function and the quasi-momentum distribution is given by

$$S(p) = \sum_{s,l=1}^k e^{ip(x_s - x_l)} \langle \hat{b}_s^\dagger \hat{b}_l \rangle. \quad (39)$$

As we do not assume the state to be translationally invariant, Eq. (39) cannot be inverted to get the two-point correlation functions in real space but we can get for integer l

$$\sum_{s=1}^k \langle \hat{b}_s^\dagger \hat{b}_{s+l} \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} dp e^{ipl} S(p), \quad (40)$$

where we have set the lattice spacing to one. If the atoms are bosons, the local Hilbert space is infinite-dimensional. However, as the interaction between the atoms must always be repulsive to ensure stability of the quantum gas, one can neglect state where more than a given cut-off number of atoms are present on a single lattice site. This allows us to work with a finite-dimensional Hilbert space. We note that this does not even need to be an approximation as one can set the maximal number of bosons per site N_S to their total number N . However, in any practical setting, one would use $N_S \ll N$ and still get a very good approximation. For an arbitrary i , we set

$$w_0^{(i)} \propto \sum_{j=1}^k (\hat{b}_i^\dagger \hat{b}_{i+j} + \hat{b}_{i+j}^\dagger \hat{b}_j) \quad (41)$$

which is traceless. Due to the sum, which stems from the absence of translational invariance, Eq. (41) is not exactly of the form given by (23) but a sum of few, i.e. logarithmically many in the Hilbert space dimension d , terms of this form. This implies

$$\|w_0^{(i)}\|_\infty \leq C \sum_{j=1}^k \|\hat{b}_j^\dagger \hat{b}_{i+j} + \hat{b}_{i+j}^\dagger \hat{b}_j\|_\infty \leq \tilde{C}k \quad (42)$$

where $C, \tilde{C} > 0$ are constants. Because $k = \Theta(\log d)$, we can employ Theorem 1 to show that a measurement of the momentum space distribution, together with an approximate 2-design, allows for efficient compressed sensing.

Although already a single choice of i yields an approximate tight frame, we can use the data corresponding to all $i = 1, \dots, k$, as they are measured anyway, to reduce the necessary number of experiments.

B. Realization of the random circuit

We now discuss how a probability measure on the nearest-neighbor unitaries that is universal can be obtained. To be as specific and simple as possible, we use the single-band Bose Hubbard model with Hamiltonian [31]

$$\hat{H} = - \sum_{i=1}^k J_i (\hat{b}_i^\dagger \hat{b}_{i+1} + \hat{b}_{i+1}^\dagger \hat{b}_i) + \frac{U_i}{2} \hat{n}_i (\hat{n}_i - 1) + \Delta_i \hat{n}_i, \quad (43)$$

where $\hat{n}_i = \hat{b}_i^\dagger \hat{b}_i$ and $J_1, \dots, J_k; \Delta_1, \dots, \Delta_k; U_1, \dots, U_k \in \mathbb{R}$. To realize the parallel random quantum circuit, we make use of the techniques of super-lattices [15, 16, 27] and speckle patterns [13]. As the total number of atoms is conserved and super-selection rules forbid the superposition of states with

different particle numbers, we restrict ourselves to particle-number conserving operations.

By using an additional lattice for which the lattice constant is twice as large, one can change the height of the wells between alternating pairs of site. This mainly affects the hopping constants J_i and to much less extent the interaction parameter U_i , an effect which we neglect. By choosing the depth of the super-lattice large enough, we get $J_i = 0$ for the non-coupled pairs and $J_i = J$ for the coupled ones. Such a double-well structure has been used in Ref. [16] to probe correlation functions. A speckle pattern is created by illuminating an uneven surface with a laser and can be modelled by a spatially fluctuating Δ_i . The situation is sketched in Figure 1. For reasons of simplicity, we only consider the regime for which the strength of the speckle potential is not correlated over the lattice sites, i.e., all Δ_i are independently distributed. Thus, we have a random circuit as in Theorem 4 and we only need to show that the corresponding gates generate a dense set in $SU_N(d_l^2)$. The local gates are

$$U_{i,j}(\Delta_1, \Delta_2, t) = \exp(-it(-J(\hat{b}_1^\dagger \hat{b}_2 + \hat{b}_2^\dagger \hat{b}_1) + \frac{U}{2}(\hat{n}_1(\hat{n}_1 - 1) + \hat{n}_2(\hat{n}_2 - 1)) + \Delta_1 \hat{n}_1 + \Delta_2 \hat{n}_2)), \quad (44)$$

where $t \geq 0$ is the time after which the super-lattice is switched and a new realization of the laser speckle is created. We assume $\Delta_{1,2}$ to be Gaussian distributed, and we have neglected global phases. Now one can adopt an argument used in Ref. [32] for showing that the Gaussian operations together with a single non-Gaussian one allow for continuous-variable quantum computation. For universality to hold, one has to generate all operations where the corresponding Hamiltonian is a polynomial in the creation and annihilation operators where every monomial must contain an equal number of creation and annihilation operators, i.e., must be balanced, to ensure particle-number conservation. This is true as Eq. (44) contains all quadratic terms and a single quartic one. Since one can generate the entire algebra generated from the original set of Hamiltonian by commutation, one can approximate an arbitrary unitary [32]. Thus, by varying $\Delta_{1,2}$, we can approximate any gate to arbitrary accuracy which implies, by continuity of Eq. (44) universality as in Definition 5. By appropriately choosing the distribution from which t is chosen, the set can be made closed under Hermitian conjugation. Now, we can apply Theorem 4 which shows together with Theorem 1 that one can perform efficient compressed sensing by using a optical super-lattices, laser speckles, and time-of-flight imaging. Again, no single site addressing is necessary, and still complete tomographic information is obtained.

VI. MORE EFFICIENT TOMOGRAPHY SCHEME FOR MATRIX PRODUCT STATES AND OPERATORS

Even though compressed sensing notably reduces the number of necessary measurements, it still scales exponentially with the number of lattice sites. Without any further assumption on the state, this cannot be overcome. However, when

the state is described by a generic matrix-product state (MPS) with a fixed bond dimension, tomography is possible with the number of measurements growing, in general, almost linearly with the system size and being constant for translationally invariant MPS. Ground states of gapped Hamiltonians which are a sum of terms acting only on a constant number of lattice sites are, generically, of this type [33].

A. Reconstructing reduced density matrices

The exponential reduction of the necessary number of measurements if the state is a MPS is due to the fact that such states are completely determined by their reduced density matrices on all blocks of l consecutive sites where l only depends on the bond dimension [9, 33]. In Ref. [9], an efficient algorithm is given for finding the MPS matrices from these reduced density matrices. Note that this procedure only works if an upper bound to the bond-dimension, or, equivalently, to the locality size of the Hamiltonian is known. This is obvious as there could always be long range correlations which do not affect the l site reduced density matrices.

Let $d_B = d_l^l$ be the dimension of the subsystem under consideration and define the operator $\mathcal{T}_{R_q} : \mathcal{B}(\mathbb{C}^d) \mapsto \mathcal{B}(\mathbb{C}^{d_B})$ be the operator acting as $\mathcal{T}_{R_q}(\rho) = \text{Tr}_{R_q} \rho$ and R_q denote all lattice sites but $q, \dots, q + l - 1$. As we only want to perform tomography on the lattice sites $q, \dots, q + l - 1$, we can trace out the remainder of the system. In this case, the tight-frame condition of Definition 2 becomes

$$\mathcal{T}_{R_q} \circ \mathcal{W}_{d_B} = \mathcal{T}_{R_q}. \quad (45)$$

If additional constraints apply, the condition reads

$$\mathcal{T}_{R_q} \circ \mathcal{W}_{d_{BN}} \circ \mathcal{P}_N = \mathcal{T}_{R_q} \circ \mathcal{P}_N \quad (46)$$

where d_{BN} is the dimension of the matrix block corresponding to the eigenvalue N of \hat{N} , restricted to the subsystem of l lattice sites. We concentrate on the former case as the second follows in an analogous way and show that a random circuit with a depth which does not depend on the system size can realize such a reduced tight frame: Assume w_0 to be a sum of traceless observables, each acting non-trivially only on some block of l lattice sites. For compressed sensing to be possible, it is sufficient for ν to induce, for every q , an approximate reduced tight frame with

$$\|\mathcal{T}_{R_q}(\mathcal{W}_{d_B, \nu} - \mathbb{1})\| \leq \sqrt{d_B(d_B^2 - 1)}\varepsilon \quad (47)$$

where ε must be chosen such that $\sqrt{d_B(d_B^2 - 1)}\varepsilon < 1/(8\sqrt{r})$, c.f. Eq. (20). This is the case if it is an approximate reduced unitary 2-design with

$$\forall q : \sup_{X, \|X\|_2=1} \left\| \text{Tr}_{R_q}(\mathcal{G}_\nu - \mathcal{G}_{H_l}^{(q)})(X) \right\|_2 \leq \varepsilon. \quad (48)$$

where $\mathcal{G}_{H_l}^{(q)}$ denotes the channel which acts as \mathcal{G}_H on a block of l sites starting at q and as the identity on the rest of the system.

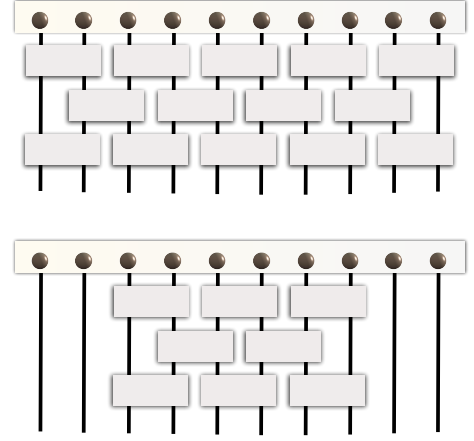


FIG. 2. Top: Parallel random circuit acting on the entire system. Bottom: Random circuit used in the proof of Theorem 6 which acts only on the inner sites. The above circuit randomizes the states of the inner sites not less than the above one.

To see that this is true, we calculate

$$\|\mathcal{T}_{R_q}(\mathcal{W}_{d_B, \nu} - \mathbb{1})\| = \sup_{X, \|X\|_2=1} \|\text{Tr}_{R_q}(\mathcal{W}_{d_B, \nu}(X) - \mathcal{W}_{d_B, H}(X))\|_2, \quad (49)$$

which yields the desired result after inserting (17) and applying (48).

To show that obtaining such a tight frame is efficiently possible, we adapt Theorem 4:

Theorem 6 (Reduced tight frames by reduced unitary 2-designs). *Let the parallel random circuit be as in Theorem 4. There exists some constant C such that it forms an ε -approximate reduced l -site 2-design as defined in (47) after $n = C \log(1/\varepsilon)l \log l$ steps.*

Proof: We show the fast convergence of our random circuit to a reduced unitary 2-design by comparing it with another random circuit which is easier to deal with, see Fig. 2. In analogy to Eq. (24) we denote by $\mathcal{G}_{\nu_l}^{(q)}$ the channel corresponding to an application of the parallel quantum circuit to a block consisting of l lattice sites starting from q . Theorem 4 implies

$$\sup_{X, \|X\|_2=1} \|\text{Tr}_{R_q}((\mathcal{G}_{\nu_l}^{(q)})^n - \mathcal{G}_{H_l}^{(q)})(X)\|_2 \leq \varepsilon \quad (50)$$

for $n = C \log(1/\varepsilon)l \log l$ where C is a constant. Using that $\lambda_2(G_{\nu_k}) \leq \lambda_2(G_{\nu_l})$, one obtains (48), concluding the proof. ■

B. Complexity of classical post-processing

Theorem 6 implies that a random quantum circuit of constant depth is sufficient to perform tomography on a reduced density matrix of constant size. The number of operations, i.e.

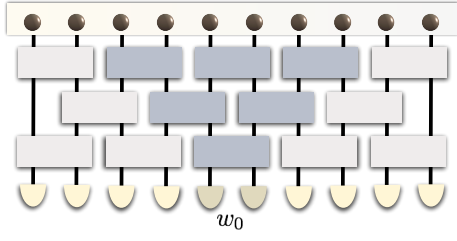


FIG. 3. Influence region for a local observables. Only the action of the darkly colored gates influences the measured observables.

random unitary gates and measurements of expectation values of w_0 , does only scale polynomially with the number of lattice sites. Therefore, we regard the quantum part of the protocol as efficient. However, this says nothing about the amount of post-processing needed because the reconstruction of the l -site reduced density matrices requires the knowledge of $\text{Tr}_{R_q} w$ for all observables w obtained by the realizations of the random quantum circuit. If one needed to keep track of the evolution of observables on the entire Hilbert space this would require an exponential amount of computational resources. To see that this is not a problem in the present situation, we use the fact that w_0 is a sum of terms which act non-trivially only on blocks of constant size. Thus, the observables induced by the constant-depth random circuit can be written as

$$w(U) = \sum_i w_i(U) \otimes \mathbb{1}_{R_i} / (d_{R_i})^{1/2} \quad (51)$$

where $w_i(U)$ acts non-trivially only on a block of constant size L starting with site i and where R_i denote all lattice sites but $i, \dots, i + L - 1$, see Figure 3. From Eq. (51), we get

$$\text{Tr}_{R_q}(w(U)) = \sum_i \sqrt{d_{R_i}} \text{Tr}_{R_q}(w_i(U)), \quad (52)$$

which means that one only needs to deal with observables on k Hilbert spaces which all have dimension d_l^L not depending on k making also the classical part of the protocol efficient. If the system is assume to be translationally invariant, all reduced density-matrices are equal, reducing the necessary number of measurements to a constant. Roughly speaking, the random circuit transforms a local observable to a reduced tight frame on l lattice sites with some influence on $L > l$ sites and none on the rest of the system.

C. Mixed states

Even though the method developed in this section is, in its present form, limited to pure states, it can be naturally extended to mixed states by using recent results on tomography

for matrix-product operators (MPO) which are a natural generalization of MPS [34, 35]. In Ref. [36], it is shown that large classes of MPO states can be efficiently reconstructed from the reduced density matrices on a constant number of sites not depending on the system size. As we have presented a method of recovering these objects, one directly obtains a way of performing tomography on states in optical lattices which are described by MPO while requiring the same experimental techniques.

VII. CONCLUSION

In this article, we have presented a new route towards efficient quantum state tomography for quantum many-body systems, specifically for bosons in optical lattices. By using random circuits, which can be implemented by means of superlattices and laser speckles, one can avoid the use of tomographically complete local measurements on single sites (in fact any local addressing) – still complete tomographic knowledge can be achieved. These are challenging, needless to say, but rely solely on time-of-flight imaging techniques, which are nowadays routinely implemented. Without any further assumptions to the state, the number of necessary measurements is optimal up to constants and logarithmic factors in the systems dimension. Restricting the set of possible states to matrix product states, both the number of measurements and the depth of the required random quantum circuit does not at all depend on the system size. This idea gives rise to the exciting perspective of actually reliably measuring out the full quantum state of a quantum many-body system in the laboratory.

There are a number of questions arising from this: For example, it would be very interesting to compare the performance of the present scheme, which is based on random circuits, with one where the quantum gates are chosen in an optimal way from some set of feasible operations. While there is not much room for improvements concerning the asymptotic behavior, the performance for small systems might differ notably. This will be studied extensively by numerical means in forthcoming work. We hope that this work stimulates further work – both of theoretical and especially experimental kind – in “quantum system identification”, in order to innovate ways of rendering quantum state tomography feasible for large quantum systems.

VIII. ACKNOWLEDGMENTS.

We thank M. Kliesch and C. Gogolin for valuable discussions, especially in the initial phase of this project. We would like to thank the EU (Qessence, Compas, Minos), the BMBF (QuOREP), and the EURYI for support.

- [3] G. M. D'Ariano and P. Lo Presti, Phys. Rev. Lett. **86**, 4195 (2001).
- [4] E. J. Candès and T. Tao, IEEE Trans. Inf. Th. **56**, 2053 (2010).
- [5] D. Gross, Y.-K. Liu, S. T. Flammia, S. Becker, and J. Eisert, Phys. Rev. Lett. **105**, 150401 (2010).
- [6] D. Gross, IEEE Trans. on Inf. Th. **57**, 1548 (2011).
- [7] A. Shabani et al., Phys. Rev. Lett. **106**, 100401 (2011).
- [8] M. Ohliger, V. Nesme, D. Gross, Y.-K. Liu, and J. Eisert, arXiv:1111.0853.
- [9] M. Cramer, M. B. Plenio, S. T. Flammia, D. Gross, S. D. Bartlett, R. Somma, O. Landon-Cardinal, Y.-K. Liu, and D. Poulin, Nat. Commun. **1**, 149 (2010).
- [10] F. G. S. L. Brandão and M. Horodecki, arXiv:1010.3654.
- [11] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, arXiv:1208.0692.
- [12] R. A. Low, PhD thesis, University of Bristol (2009).
- [13] D. Clément, A. F. Varón, J. A. Retter, L. Sanchez-Palencia, A. Aspect, and P. Bouye, New J. Phys. **8**, 165 (2006).
- [14] D. A. Abanin and E. Demler, Phys. Rev. Lett. **109**, 020504 (2012); A. J. Daley, H. Pichler, J. Schachenmayer, and P. Zoller, Phys. Rev. Lett. **109**, 020505 (2012).
- [15] S. Trotzky, Y.-A. Chen, U. Schnorrberger, P. Cheinet, and I. Bloch, Phys. Rev. Lett. **105**, 265303 (2010); G.-B. Jo, J. Guzman, C. K. Thomas, P. Hosur, A. Vishwanath, and D. M. Stamper-Kurn, Phys. Rev. Lett. **108**, 045305 (2012).
- [16] S. Trotzky, Y.-A. Chen, A. Flesch, I. P. McCulloch, U. Schollwöck, J. Eisert, and I. Bloch, Nature Physics **8**, 325 (2012).
- [17] A. W. Harrow and R. A. Low, Commun. Math. Phys. **291**, 257 (2009).
- [18] C. Dankert, R. Cleve, J. Emerson, and E. Livine, arXiv:quant-ph/0606161.
- [19] D. Gross, K. M. R. Audenaert, and J. Eisert, J. Math. Phys. **48**, 052104 (2007).
- [20] E. J. Candès and Y. Plan, arXiv:1011.3854.
- [21] S. T. Merkel, C. A. Riofrio, S. T. Flammia, and I. H. Deutsch, Phys. Rev. A **81**, 032126 (2010).
- [22] D. Burgarth and K. Yuasa, Phys. Rev. Lett. **108**, 080502 (2012).
- [23] B. Collins and P. Sniady, Comm. Math. Phys. **264**, 773 (2006).
- [24] M. Greiner, O. Mandel, T. Esslinger, T. W. Hänsch, and I. Bloch, Nature **415**, 39 (2002).
- [25] R. Jördens, N. Strohmaier, K. Günter, H. Moritz, and T. Esslinger, Nature **455**, 204 (2008).
- [26] M. Lewenstein, A. Sanpera, V. Ahufinger, B. Damski, A. Sen, and U. Sen, Adv. Phys. **2**, 243 (2007); I. Bloch, J. Dalibard, and S. Nascimbene, Nature Physics **8**, 267 (2012); J. I. Cirac and P. Zoller, Nature Physics **8**, 264 (2012).
- [27] P. Barmettler, A. M. Rey, E. Demler, M. D. Lukin, I. Bloch, and V. Gritsev, Phys. Rev. A **78**, 012330 (2008); A. Flesch, M. Cramer, I. P. McCulloch, U. Schollwoeck, and J. Eisert, Phys. Rev. A **78**, 033608 (2008).
- [28] R. Bhatia, *Matrix analysis*, Springer (1997).
- [29] J. Eisert, M. Cramer, and M. B. Plenio, Rev. Mod. Phys. **82**, 277 (2010).
- [30] A. Hoffmann and A. Pelster, Phys. Rev. A **79**, 053623 (2009).
- [31] M. P. A. Fisher, P. B. Weichman, G. Grinstein, and D. S. Fisher, Phys. Rev. B **40**, 546 (1989).
- [32] S. Lloyd and S. L. Braunstein, Phys. Rev. Lett. **82**, 1784 (1999).
- [33] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Quant. Inf. Comp. **7**, 401 (2007).
- [34] F. Verstraete, J. J. Garcia-Ripoll, and J. I. Cirac, Phys. Rev. Lett. **93**, 207204 (2004).
- [35] M. Zwolek and G. Vidal, Phys. Rev. Lett. **93**, 207205 (2004).
- [36] T. Baumgratz, D. Gross, M. Cramer, and M. B. Plenio, arXiv:1207.0358.